AFRL-RH-WP-TR-2013-0017

CO-ADAPTIVE AIDING AND AUTOMATION ENHANCE OPERATOR PERFORMANCE

James C. Christensen & Justin R. Estepp
Air Force Research Laboratory, Wright-Patterson Air Force Base, OH 45433

March 2013

FINAL REPORT

Distribution A: Approved for public release; distribution unlimited.

AIR FORCE RESEARCH LABORATORY
711 HUMAN PERFORMANCE WING,
HUMAN EFFECTIVENESS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR-2013-0017 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//**signature//**
James Christensen
Program Manager
Collaborative Interfaces Branch

//**signature//**
William E. Russell
Chief, Collaborative Interfaces Branch
Warfighter Interface Division

//**signature//**
Michael A. Vidulich
Technical Advisor
Warfighter Interface Division
Human Effectiveness Directorate
711 Human Performance Wing

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS**.

1. REPORT DATE *(DD-MM-YY)*
01/03/13

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
2 JAN 2010 – 16 APR 2013

4. TITLE AND SUBTITLE
Co-adaptive aiding and automation enhance operator performance

5a. CONTRACT NUMBER
IN-HOUSE

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
James C. Christensen & Justin R. Estepp

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
(H041) 71840879

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Applied Neuroscience Branch
Applied Adaptive Aiding Section
Wright-Patterson AFB, OH 45433-7951

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711 Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Applied Neuroscience Branch
Wright-Patterson Air Force Base, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711HPW/RHCPA

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2013-0017

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
88 ABW Cleared 05/13/2013; 88ABW-2013-2274. Report contains color.

14. ABSTRACT
In this work, we demonstrate performance improvements over one week with a physiologically activated adaptive aiding system in a simulated remotely piloted aircraft control task. No such improvement occurred with manual activation, leading to the conclusion that operators adapted their physiological signals to better manage the aiding, hence co-adaptation.

15. SUBJECT TERMS
neuroergonomics, mental workload, function allocation, human/computer interaction

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
Unclassified	Unclassified	Unclassified	SAR	25	James C. Christensen **19b. TELEPHONE NUMBER** *(Include Area Code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

TABLE OF CONTENTS

Section **Page**

LIST OF FIGURES

Page

ACKNOWLEDGEMENTS

The authors wish to thank Iris Davis, Justin Cheuvront, and Margaret Bowers for their considerable efforts in implementing and running this study. The authors would also like to thank Dr. Glenn Wilson for providing insightful comments on the design and analysis, and Dr. Joel Warm for many comments that have improved this manuscript. Lastly, we would like to thank Ms. Krystal Thomas for her assistance in preparing this manuscript. Funding for this work was provided by the Air Force Office of Scientific Research.

1.0 SUMMARY

This report covers progress made in real-time monitoring and adaptive aiding to increase operator performance. Progress from 2008 to 2010 has been previously reported in our published work (Christensen et al 2010; Christensen, Estepp, Wilson & Russell 2012). This report covers progress made from 2010 to the present. The present work expands upon the theory of adaptive aiding by measuring the effectiveness of co-adaptive aiding, wherein we explicitly allow for both system and user to adapt to each other. Adaptive aiding driven by psychophysiological monitoring has been demonstrated to be a highly effective means of controlling task allocation and system functioning. Psychophysiological monitoring is uniquely well-suited for co-adaptation, as malleable brain activity may be used as a continuous input to the adaptive system. To establish the efficacy of the co-adaptive system, physiological activation of adaptation was directly compared with manual activation or no activation of the same automation and cuing systems. We used interface adaptations and automation that are plausible for real-world operations, presented in a multi-Remotely Piloted Aircraft (RPA) control simulation. Participants completed three days of testing over one week. Performance was assessed via proportion of targets successfully engaged. In the first two days of testing, there were no significant differences in performance between the conditions. In the third session, physiological adaptation produced the highest performance. By extending the data collection over multiple days, we offered enough time and repeated experience for user adaptation and online system adaptation, hence demonstrating co-adaptive aiding.

2.0 INTRODUCTION

Adaptive aiding and adaptive automation have been studied and discussed for quite some time. The idea of reallocating tasks between human and machine (Rouse, 1976) to produce maximum overall system performance has an intuitive appeal that has led to great interest over the years. There are many possible methods of determining task allocation in an adaptive system, but one that has shown particular promise is based on psychophysiological monitoring (Byrne & Parasuraman, 1996; Parasuraman & Wilson, 2008; Pope, Bogart, & Bartolome, 1995; Scerbo, 2007) and real-time modification (Wilson & Russell, 2003; Wilson & Russell, 2007). The present study was intended to build upon this work in two key areas: increasing the realism of the adaptive aiding used and extending data collection over multiple sessions and days. This extension over days allows time for the operator to adapt to the system as well as adapting the system to the moment-by-moment capabilities of the operator, demonstrating co-adaptation. In order to strongly test for improvements with physiologically triggered aiding, physiologically triggered adaptive aiding was compared to manually triggered adaptable aiding, based on the work of Bailey, Scerbo, Freeman, Mikulka, and Scott (2006).

Wilson and Russell (2007) presented a particularly effective demonstration of the potential of real-time, physiologically activated adaptive aiding. They used a complex simulation of operating multiple Remotely Piloted Aircraft (RPA) in a ground attack mission while monitoring electroencephalographic (EEG), electrocardiographic (ECG) and electrooculographic (EOG) data (hereafter referred to collectively as physiological data). The physiological data were classified in real time by trained artificial neural networks. The

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

neural networks continuously classified operator workload and initiated aiding when high workload was detected. The aiding consisted of reducing the airspeed of those vehicles that were approaching a target by 50%. They demonstrated that this aiding resulted in significantly improved operator performance when activation was based on the physiological workload rather than random activation. This provided a strong demonstration of the potential positive impact of real-time activation of aiding based on physiologically assessed workload.

Reducing the airspeed by half and consequently reducing the rate at which events had to be processed is clearly a highly effective means of mitigating task demands on the operator. However, in many application settings this type of task slowing may not be a realistic option. Extensively developing effective and realistic demand reduction techniques was not, in fact, one of Wilson and Russell's (2007) objectives, as the primary goal was to demonstrate the potential of such adaptive aiding systems. To further advance this area of work, it is necessary to demonstrate the effectiveness of more realistic aiding techniques appropriate to multi-RPA operation. To address this, the present study used an adaptive system that implemented a suite of mitigation procedures corresponding to stages two and three in the hierarchy proposed by Fuchs, Berka, Levendowski, and Juhnke (2006): directing attention via cuing and enhancing the salience of critical events via decluttering. In addition, limited automation was invoked that paired vehicles and targets, subject to operator override. This builds on a similar study performed by Parasuraman, Cosenzo, and De Visser (2009) by testing attentional aiding in addition to automation.

One of the key potential benefits offered by physiological activation is that the system should have little or no workload associated with managing activation (Byrne & Parasuraman, 1996), unlike manually activated aiding. However, no physiologically based workload monitoring system has yet achieved perfect detection of high workload due, at least in part, to noise or artifacts (Smith, Gevins, Brown, Karnik, & Du, 2001). Consequently, such systems exhibit errors and activate aiding when it is unnecessary or possibly detrimental to performance. While the operator could make similar errors in manually activating the aiding system, a critical question is whether overall performance achievable with an imperfect physiological activation system exceeds that possible with manual activation. Undoubtedly, the answer depends on the accuracy of physiologically-driven activations as well as the workload involved in manual management of activation. Bailey, Scerbo, Freeman, Mikulka, and Scott (2006) demonstrated modest performance improvements for physiologically activated automation as compared to manually activated; this study builds upon that work by allowing significantly greater time for participants to adapt to the physiological activation system.

There is a close relation between physiologically activated adaptive aiding and brain-computer interfaces (BCI). BCI here refers to the use of brain signals to directly control systems, such as the classic example of communicating via selection of letters based on analysis or classification of EEG signals (Farwell & Donchin, 1988). Physiologically activated adaptive aiding is, in a sense, a special case of BCI wherein the purpose is not direct control but rather monitoring and providing aiding to the operator to enable them to work more effectively (often referred to as passive BCI, e.g. Zander, Kothe, Jatzev, &

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

Gaertner, 2010). Researchers investigating BCI have noted that brain signals used for BCI applications generally evolve with feedback in a manner that improves system performance. Recognizing that brain signals are adapting in BCI paradigms suggests that similar ongoing adaptation of analysis techniques or classifiers (e.g., Wolpaw, McFarland, Neat, & Forneris, 1991) is a promising approach. However, BCI work demonstrating the value of recognizing that both the operator and physiologic classification systems are adaptive controllers has only been accomplished recently (Huang, Erdogmus, Pavel, Mathan, & Hild II, 2011; McFarland, Sarnacki, & Wolpaw, 2011). In this line of work, operators that were trained to use a BCI system over a period of days achieved greater accuracy with practice. Analysis revealed that operators adapted their own EEG signals to more precisely trigger the system. This mutually adaptive relationship in BCI (as proposed by Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002) may then be extended to application in a physiologically activated adaptive interface. In an adaptive interface, easily visible interface changes provide the necessary feedback to operators who may then adapt brain activity to improve their management of the interface. Conversely, the physiological activation system may be adapted to the operator by retraining classifiers to capitalize on changes in user brain activity. This study retrained classifiers for each participant and day in order to adapt to such changes. We term this application and overall process co-adaptation, hence the present study tested an implementation of a co-adaptive interface.

The extension of data collection to multiple days was motivated by previous results in which BCI-adapted EEG was shown to induce neuroplasticity (Ros, Munneke, Ruge, Gruzelier, & Rothwell, 2010), and separate work which has demonstrated that neural activity during sleep is a key factor in plasticity and the consolidation of learned items (Steriade, 2001). The first and second days of collection in the current study cover a normal sleep-wake cycle, and were chosen as most likely to support adaptive changes in the operator’s neural signals. The third day of data collection was spaced to one week after the first; this was based on learning research that has demonstrated that expanding the inter-study interval can lead to more stable retention (Hser & Wickens, 1989), as well as best accommodating participant availability.

The present study therefore sought to compare performance in a realistic task with a physiologically-activated adaptive aiding system to the same system when triggered by simple manual keypress. The study sessions were conducted over three days in order to facilitate neuroplastic changes associated with normal sleep-wake cycles and expanding inter-study intervals. By thus supporting user adaptation to an already-adaptive system, the study ultimately facilitated co-adaptation.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

3.0 METHODS

Ten[1] naïve persons either currently employed at Wright-Patterson Air Force Base or students at Wright State University in Dayton, Ohio, volunteered to participate in the study. Employees of Wright Patterson Air Force Base received their normal duty pay, and Wright State students were paid $15 per hour for their participation. Participants completed comprehensive written informed consent prior to the start of the experiment. All reported normal or corrected-to normal vision with no color blindness. Two participants were female and eight male, all between the ages of 19 and 26 years (mean of 22 years). All study procedures were reviewed and approved by the Air Force Research Laboratory Institutional Review Board.

The multi-RPA operation task was a PC-based supervisory control simulation of a "suppression of enemy air defense" mission (Schmidt, Wilson, Funke, Davis, & Caldwell, 2010) using keyboard and mouse controls. Participants monitored the progress of eight or 16 generic RPAs on two abutted 51 cm (diagonal) computer screens as they flew a preplanned mission. When the RPAs came within radar range of a target, simulated radar images of the target area were automatically acquired and operators could then mouse-click a map icon to download and view these images. Each image contained zero to eight targets drawn from three visually distinct types, as well as 25 to 30 nontarget distractors. Each target type was to be engaged with a specific weapon type, generically termed small, medium, and large. Each RPA carried a limited number of two of the three weapon types. After visually searching the images, participants were required to select each weapon with the mouse and click the image to mark appropriate weapon-target pairings before the RPA reached the minimum weapons release distance. These tasks were performed for each of the RPAs as targets came in range. If the targets were not selected and/or the weapons release command (mouse click on a confirmation box) was not given in time, the weapons from that RPA could not be released thereby reducing the number of targets successfully engaged. Misidentifying a target and assigning an inappropriate weapon type was not counted as success. Participants could use the mouse to designate waypoints and direct RPAs away from pre-planned routes but were not allowed to double back to reengage targets. Successfully engaging all targets required rerouting RPAs as stores were expended. This task was intended to broadly represent future operator control tasks and tap a wide range of cognitive skills such as working memory, visual search, object recognition, task switching, and flexibly managing conflicting priorities.

Adaptive aiding and automation were implemented in this task via several methods. The aiding methods were intended to cue attention to time-critical tasks. Because the aircraft could not double back, a simple time to contact calculation enabled prioritization of those targets that most urgently needed operator attention. The top three target-RPA pairs were then color coded with red, yellow, and green transparent circles. All other RPAs and targets had contrast lowered ("fog layer") to render them less salient (Fig. 1). In addition to these

---

[1] Six additional participants (five male and one female) were enrolled in the study. The total duration of the study including training and testing typically extended over several months. This resulted in these participants leaving prior to completion, most commonly due to moving or graduation.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

aiding techniques, the interface supported partial task automation. Instead of the operator manually assigning a RPA to a target, the automation linked each target to the closest RPA. This linking was made without regard to the weapon type required or stores on that RPA, and thus could link RPAs that were incapable of engaging a target. The operator could see the stores on each RPA and override incorrect links. Lastly, when a target came into range the automation displayed the appropriate simulated radar image rather than requiring the operator to request it. All of the aiding techniques and automation were activated and deactivated together, either on operator command or based on physiological workload classification. Operator command (the manual activation condition) consisted of pressing the spacebar on the keyboard. Physiological activation was triggered based on the output of a physiological workload classifier that replicated Wilson and Russell (2007).

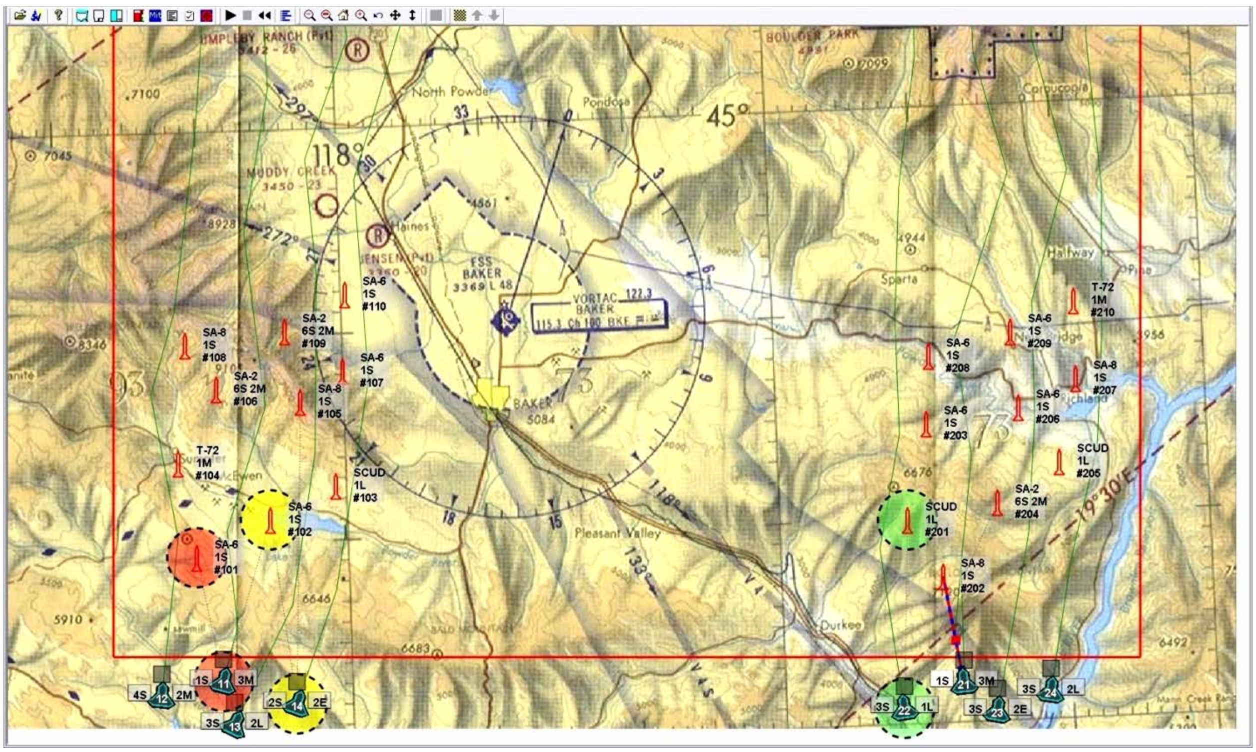

Figure 1. An example of the multi-RPA control task. The wedge shaped symbols at the bottom represent RPAs that the operator is controlling; the missile-shaped symbols represent air defense sites that need to be engaged. Aircraft proceeded from south to north at a fixed rate, requiring approximately 20 minutes to traverse the target area. The color cuing is active in this example, and is visible as same-color circles around a target/RPA pair.

One concern with the workload classification based aiding is that effective aiding should reduce workload, leading to deactivation of the aiding. This could then lead to a high frequency of activation/deactivation as workload is modulated by the aiding. In order to reduce this, any change in aiding (either manually activated or physiologically activated) triggered a timeout of 15 seconds before allowing another such change. Aiding remained on if reactivation was commanded during this 15 second window.

Physiological data recording and analysis replicated Wilson and Russell (2007). Briefly, this included EEG from five channels, at F7, Fz, Pz, T5, and O2, positioned according to the

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

International 10-20 electrode system (Jasper, 1958), using an Electro-Cap (Electro-Cap International, Inc., Eaton, OH). Reference and ground electrodes were positioned on the mastoid processes, with impedances verified below 5 kiloohms. Horizontal and vertical electrooculogram (HEOG and VEOG, respectively) and electrocardiogram (ECG) were also recorded using standard Ag/AgCl electrodes. A Cleveland Medical Devices, Inc. (now Great Lakes NeuroTechnologies, Cleveland, OH) BioRadio 110 telemetry unit was used to acquire these data channels with a sampling rate of 200 Hz (12 bit resolution, bandpass filtered between 0.5 and 52.4 Hz). Corrections for eye movement and blinks were made using an on-line implementation of an adaptive filter with HEOG and VEOG used as reference noise channels (He, Wilson & Russell, 2004; He, Wilson, Russell & Gerschutz, 2007). Inter-beat interval (IBI), calculated over a ten second window, and was derived from the ECG channel using an on-line algorithm (Pan & Tompkins, 1985; Hamilton & Tompkins, 1986). Similarly, blink rate, calculated over a thirty second window, was derived using an on-line algorithm developed by Kong & Wilson (1998). The EEG data were filtered into separate band-limited channels using elliptical IIR filter banks. The passbands for each channel were consistent with the five traditional bands of EEG: delta (0.5-3 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz) and gamma (31-42 Hz). For real-time classification, the data were segmented into ten-second windows with a nine-second overlap. Log power of the five bands from the five sites was used in addition to IBI and blink rate, resulting in 37 features as inputs for workload classification. Workload classification was accomplished via a feedforward artificial neural network (ANN) trained via backpropagation with two output nodes corresponding to low and high workload, replicating the procedure used by Wilson and Russell (2003). Immediately following each trial, participants completed the computerized NASA-TLX (Hart & Staveland, 1988).

Participants were extensively trained on the task prior to the experimental sessions in order to reduce training effects during the three data collection days. Each participant completed a minimum of two hours of training per day for five days. Over the course of these days of training, participants were gradually introduced to all of the task features. The aiding system was introduced on approximately the fourth day of training; the fourth and fifth days included practice manually controlling the aiding system. Participants were instructed to maximize their performance by using aiding as much or as little as they liked when it was available. Training was continued until a performance criterion (80% of targets engaged without any aiding while controlling 16 RPAs) was met.

Testing days for each participant followed a fixed schedule: Day Two was always right after Day One, with Day Three one week after Day One (six days after Day Two). This schedule was chosen in order to maximize the potential for operator adaptation in using the adaptive aiding system, space the learning sessions over an expanding interval, and maximize participant availability. On each testing day, after fitting all electrodes the participants completed one 5 minute task run to warm up. A subsequent resting baseline included two minutes of eyes open and two minutes of eyes closed. Setting up the physiologically activated aiding then required two classifier training trials, one 10 minute trial each of low difficulty (8 RPAs) and one high (16 RPAs), with no aiding. Twenty minutes of data was sufficient to train the classifier, however significantly less data (10 minutes total) has proven sufficient in other similar studies (Christensen, Estepp, Wilson, & Russell, 2012). These

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

classifier training trials were repeated on each new day of testing in order to adapt the classifier to changes that may have taken place in operator brain signals. In order to avoid order or carryover effects, we randomized the order of trials for each participant and day. However, the classifier training trials had to precede any run using physiologically activated aiding. With this constraint and three experimental trials (one each of no aiding, physiologically activated aiding and manual activated aiding) there are 72 possible randomizations. Thirty were selected that resulted in each participant completing the three aiding conditions in a unique order each day, with the physiologically activated condition appearing once in each possible position. The experimental trials were 20 minutes in duration, one each with no aiding, manually activated aiding, and physiologically activated aiding. Each experimental trial included a low difficulty middle segment of approximately 5 minutes during which only 8 of the 16 RPAs had targets to engage; the remainder of the trial was high difficulty with all 16 RPAs engaging targets. A total of 120 targets were presented in these high difficulty segments. The embedded low-difficulty segment enabled verification that the physiologically activated aiding did not simply remain on for the entirety of a run. Participants completed the NASA-TLX immediately following each trial. In total, the testing session took approximately five hours each day. Performance data will be reported for the high-difficulty (16 RPAs controlled) portions only as low-difficulty portions are at ceiling.

4.0 RESULTS

Performance was calculated as the average proportion of targets successfully engaged in the high-difficulty segments of a trial (out of 120 high-difficulty targets total per trial). For statistical analysis, these proportions were corrected via the arcsine transform (Freeman & Tukey, 1950); all data reported in figures are raw proportions to aid in interpretation. A 3 (days) x 3 (no aiding, physiologically activated, manually activated) repeated-measures ANOVA was conducted using the Huynh-Feldt correction for any violations of sphericity. The main effects of aiding activation type and day of testing were not significant, $p>.1$. However, there was a significant interaction between the two factors, $F(3.06, 27.6)=5.37$, $p=.005$, partial $\eta^2=.374$ (Fig. 2). It is evident from Figure 2 that the aiding groups did not differ on the first two days, but physiological activation improved on the third day. This interaction was probed with individual ANOVAs that compared the three aiding conditions separately for each day; for the first two days, these tests were not significant, $p>.1$. On the third day, this test was significant, $F(2,18)=5.83$, $p<.01$, partial $\eta^2=.421$. Post hoc comparisons using the Tukey HSD test indicated that the mean for the physiological activation condition (M = .90, SD = .05) was significantly greater than the means for the manual activation (M = .84, SD = .06) and no aiding conditions (M = .82, SD = .07). The latter two conditions were not significantly different from each other. The magnitude of the difference between physiological activation and manual activation on the third day was modest at 6%; however, the effect size (Cohen's *d)* for this comparison was large at 1.17.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

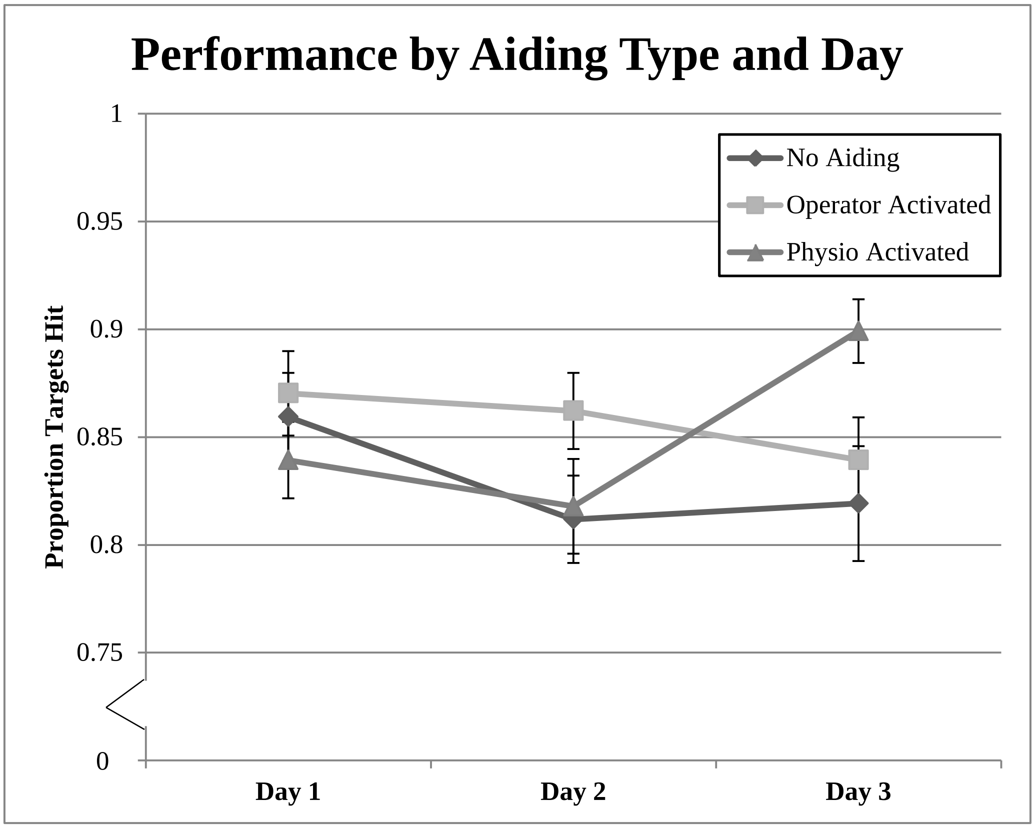

Figure 2. Performance results expressed as the mean proportion of total targets successfully engaged in the high-workload portions of a trial. "Operator" refers to operator control of aiding activation, while "Physio" refers to physiological activation of aiding. Error bars are standard errors.

The NASA-TLX subjective workload data were analyzed similarly to the performance data reported above. The overall scores were calculated via the unweighted procedure (Nygren, 1991). The task was indeed challenging; the mean overall score for high difficulty conditions was 61 as opposed to 24 for the low difficulty conditions. There was a great deal of apparently random variance in scores within participants across days, even at constant performance; to ameliorate this, scores were normalized via z-scoring within participant and day as recommended by Stevens (1971). Means and standard deviations for z-scoring were drawn from all eight composite scores (within participant and day): two from the calibration conditions, and two from each of the three aiding conditions (one for the low difficulty segment and one for the high difficulty). A possible consequence of this normalization is a reduction in between-day differences if those differences are scalar; as we did not observe a main effect of days on performance this was judged acceptable. Even after normalizing, one participant produced outlier TLX scores that were approximately four standard deviations from the mean; this participant was excluded from all subsequent analyses of these data. A 3 x 3 (three aiding conditions and three days) repeated-measures ANOVA revealed a significant main effect of day, $F(1.8,14.5)=13.01$, $p=.001$, partial $\eta^2=.619$ and a significant interaction, $F(2.8,22.3)=3.09$, $p=.035$, partial $\eta^2=.296$ (Fig. 3). As with the performance data, individual ANOVAs compared the aiding conditions separately for each of the three days. These revealed no significant differences on Day 1 or Day 2, but a significant main effect on Day 3, $F(2,16)=12.08$, $p=.001$, partial $\eta^2=.602$. Post hoc

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

comparisons using the Tukey HSD test indicated that all three aiding conditions are significantly different from each other, with no aiding producing significantly higher workload (M = 1.2, SD = .2) than the operator activated condition (M = .94, SD = .2), which in turn was higher than the physiological activation condition (M = .69, SD = .2).

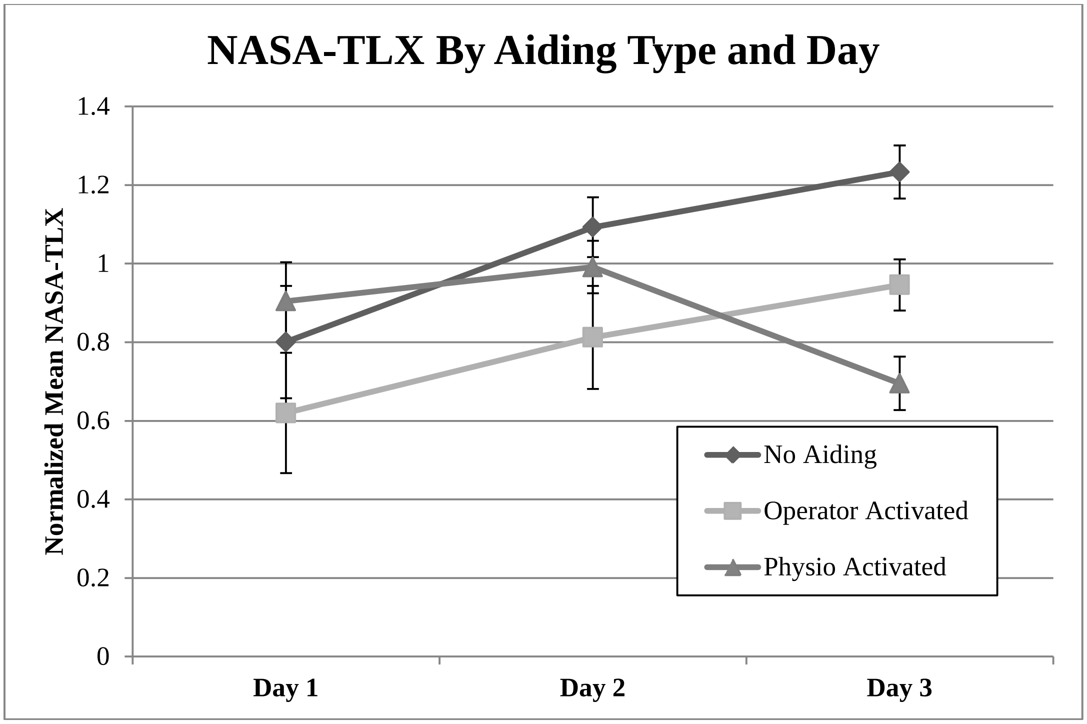

Figure 3. NASA-TLX overall workload scores expressed as the normalized mean score across participants. "Operator" again refers to operator control of aiding activation, while "Physio" refers to physiological activation of aiding. Error bars are standard errors.

As the activation of the aiding was not controlled across conditions, it was necessary to test if the difference in performance may be simply explained by a difference in the proportion of time the aiding was active. A 2 x 3 (aiding activation type by days) repeated-measures ANOVA was run on the proportion of the high workload segments of a trial during which the aiding was active. This ANOVA did not result in any significant effects; the closest to significance was the aiding-by-day interaction, $F(1,9) = 2.50$, $p>.1$. Given the significant differences between physiological activation and the other conditions on the third day reported above, the difference between operator activation and physiological activation was checked on Day Three via a simple paired t-test. This was likewise not significant, $t(9)=1.01$, $p>.3$ (two tailed). For Day Three, operators activated the aiding for an average of 49% of the run, while the physiological activation resulted in aiding for 61% of the high workload portions of a trial. This same analysis also indicated that the physiological activation system discriminated between the high and low workload segments; it was active for an average of 23% of the time during the low workload segments. Due to the variable nature of participant workload during a run (due to differences in individual strategy, RPA routing, weapon stores usage, etc.), it is difficult to assign a "correct" classification percentage for the physiological activation system. It is nonetheless intuitive that aiding should be more active in segments of increased workload.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

The 15-second timeout on activation/deactivation of aiding reduced but did not eliminate more frequent cycling of aiding in the physiologically activated condition. This was probed with a 2 x 3 (aiding activation type by days) repeated-measures ANOVA conducted on the number of activations. There was a significant main effect of activation type, $F(1,9)=35.09$, $p<.01$. Across days, aiding was triggered an average of five times per run in the manually activated aiding condition, and 15 times in the physiologically activated condition. There was no significant effect of testing day or interaction effect ($p>.1$). Aiding cycled more frequently in the physiologically activated condition but was not on for a significantly greater length of time.

5.0 DISCUSSION

This work set out to test the efficacy of simultaneously adapting an interface to the user via physiological monitoring while facilitating the possibility of neuroplastic changes associated with a normal sleep-wake cycle and inter-study intervals. The significant interaction between type of aiding and days of testing reveals that this approach indeed resulted in improved performance, above that achieved with a simple manually controlled adaptable interface, which is generally consistent with Bailey, Scerbo, Freeman, Mikulka, and Scott (2006). This is perhaps the key criterion for a successful physiological aiding system: despite imperfect workload classification it is still superior to manual control.

Performance was not significantly different over the three days in the no aiding and the manually activated aiding conditions. This validates that task training did indeed approach asymptote; if participants were still learning the task we would have expected a simple main effect of days, regardless of aiding condition. Similarly, the training with the manually activated aiding was effective; if they were still learning to use the aiding system we would expect an interaction between aiding condition and days, with significantly better performance in both of the aided conditions relative to the unaided condition. There were no significant differences between manually activated aiding and no aiding in performance; this is evidence that learning or adaptation is confined to physiologically activated aiding, consistent with our hypothesis regarding co-adaptive aiding.

Analysis of the NASA-TLX workload scores revealed that subjective workload mirrors performance to some degree; the same interaction between aiding and day was observed, although post hoc testing revealed that all three aiding conditions were different from each other on Day 3, whereas in the performance analyses operator activation and no aiding were not significantly different from each other on Day 3. The pattern of differences in TLX on Day 3 matches expectations, with no aiding producing the highest workload, operator activation the next highest, and physiological the lowest. This confirms that on Day 3, the physiological activation produced both the highest performance and the lowest subjective workload. The manually activated aiding is not completely ineffective; subjective workload was lower on Day 3 than that observed with no aiding even though there was no significant difference in performance.

The present work was not designed to fully elaborate the nature of adaptive changes occurring within the operator. However, we may reasonably infer that the process may

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

involve strategic changes in their approach to the task, changes in physiological signals that reduce noise (McFarland, Sarnacki, & Wolpaw, 2011), and perhaps the development of conscious control of signals to activate aiding due to what may be considered a biofeedback paradigm. Strategic changes would have to be uniquely effective with physiologically activated aiding to produce the observed pattern of results and were not reported by participants in post-study debriefing. On the other hand, greater facility with the physiological activation system as a result of adaptive changes in their own physiological signals is consistent with the existing literature on neuroplasticity and adaptive BCI. It is also not possible to determine to what degree co-adaptation requires conscious effort or intention to adapt; there is a large body of literature on implicit learning (Reber, 1989) that suggests conscious effort may not be required. As all such adaptive changes on the part of the operator presumably have a neural substrate, it may not be possible to separate the exact causes or sources of co-adaptation, i.e. a reduction in noise in physiological workload classification could be an ancillary effect of neural changes associated with strategic change.

While the proportion of time in which aiding was active was equivalent between the manually activated and physiologically activated aiding conditions, this was achieved with more on/off cycling in the physiologically activated condition. This is consistent with the expectation that effective aiding should reduce workload, leading to deactivation of the aiding, though simple errors in workload classification may also contribute. These factors result in more frequent transitions from both low to high workload and high to low workload; the performance effects of these transitions may have cancelled each other out (Matthews, 1986) or produced a net decline in performance (Krulewitz, Warm, & Wohl, 1975) relative to the less frequent transitions in the manually activated aiding condition. It is therefore possible that the observed improvement in performance with physiologically activated aiding would have been increased with less frequent transitions, perhaps achieved via a longer timeout between such changes. The management of workload transitions in an adaptive aiding context will require careful attention in future work.

The overall improvement in performance obtained with this realistic combination of automation and task cuing was nonetheless relatively modest. At best (Day 3, with physiological activation as compared to no aiding) the proportion of targets engaged improved from .81 to .90, or 9%. However, this improvement is similar to the 12% improvement observed in Bailey et al (2006), and the 15% improvement observed in Prinzel et al (2000), both with more artificial tasks. While this may be coincidental, it is possible that the decreased workload associated with not having to manually manage an adaptive interface in a complex task should result in these levels of modest performance improvements.

It is also of note that the manually controlled aiding resulted in performance no better than that observed with no aiding at all. By chance, this study may have happened upon a combination of aiding effectiveness, workload associated with managing the aiding, and task difficulty that resulted in the additional management workload cancelling out any benefits from having the aiding. If that is true, then the physiologically activated condition reveals the underlying effectiveness of the aiding system as it does not burden the operator

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

with additional workload to manage activation. We may still then conclude that the aiding used in this study is of relatively marginal benefit, which points at a fundamental limitation in this line of research: many or most system adaptations and automation that improve performance overall in a task are interface improvements that should be used at all times. The promise of adaptive interfaces is predicated on the assumption that there is some cost associated with constant usage, such that performance is less than optimal. For the experimenter who wishes to study adaptive aiding in realistic tasks, this creates a challenging problem in expert system and interface design: develop aiding or automation that is useful under condition of high load, but detrimental under low load. In this study, this condition was met by the use of imperfect, simple rule-based automation; a non-overloaded operator could outperform the automation and thus should not use it. As has been discussed (Fuchs, Berka, Levendowski, & Juhnke, 2006), in tasks such as RPA operations a considerable amount of state information must be available for automation to be effective. In this study, the state of the environment is known absolutely, and a reasonable approximation of ideal user behavior is possible due to the constrained nature of the task. In order to find real-world application, adaptive interface techniques must be either constrained to a very narrow window of user behavior (i.e. avoiding controlled flight into terrain as in automatic ground collision avoidance) or utilize artificial intelligence that is considerably more sophisticated than systems widely available now. Based on Wilson and Russell (2007), much better performance improvements are achievable with physiological activation of highly effective aiding techniques; advancement in this area will be required for operationally effective adaptive interfaces.

Outside operational applications, the challenge of implementing effective and appropriate aiding may be avoidable. In training or team-based applications (e.g. Elkins et al., 2009; Espevik, Helge Johnsen, Eid, & Thayer, 2006; Stevens, Galloway, Berka & Sprang, 2009), it may be sufficient to provide additional state information via physiological workload monitoring, thus enabling either a human instructor or human teammates to more effectively adapt their own behavior, or adapt the team composition more appropriately to the task (e.g. Woolley et al., 2007).

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

6.0 CONCLUSIONS

In summary, the present work demonstrated the effectiveness of realistic co-adaptive aiding in a simulated multi-RPA control task. Over time, the users adapted their interaction with physiologically activated aiding, while that interface adapted to them, hence co-adaptive aiding.

KEY POINTS

- Performance with physiologically activated adaptive aiding on the third day of testing exceeded performance with manually activated aiding or no aiding
- There was no significant effect of activation type on the first two days
- Manually controlled aiding is a key comparison for alternative control mechanisms
- Effective adaptation or mitigation techniques are difficult to design and implement in realistic tasks

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

7.0 REFERENCES

Bailey, N. R., Scerbo, M. W., Freeman, F. G., Mikulka, P. J., & Scott, L. A. (2006). Comparison of a brain-based adaptive system and a manual adaptable system for invoking automation. *Human Factors, 48* (4), 693-709.

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42*, 249-268.

Christensen, J. C., Estepp, J. R., Wilson, G. F., & Russell, C. A. (2012). The effects of day-to-day variability of physiological data on operator functional state classification. *NeuroImage, 59* (1), 57-63.

Christensen, J.C., Estepp, J.R., Davis, I.E., & Wilson, G.F. (2010) Learning effects in physiologically activated adaptive aiding. In T. Marek, W. Karwowski, & V. Rice (Eds.), *Advances in Understanding Human Performance: Neuroergonomics, Human Factors Design, and Special Populations.* Boca Raton, Florida: Taylor and Francis.

Elkins, A.N., Muth, E.R., Hoover, A.W., Walker, A.D., Carpenter, T.L., & Switzer, F.S. (2009). Physiological compliance and team performance. *Applied Ergnonomics*, *40*, 997-1003.

Espevik, R., Helge Johnsen, B., Eid, J., & Thayer, J.F. (2006). Shared mental models and operational effectiveness: Effects on performance and team processes in submarine attack teams, *Military Psychology, 18 (Suppl.)*, S23-S36.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, *70*(6), 510-523.

Freeman, M. F. & Tukey, J. W. (1950). Transformations Related to the Angular and the Square Root. *Annuals of Mathematical Statistics, 21(4)*, 607-611.

Fuchs, S., Berka, C., Levendowski, D., & Juhnke, J. (2006). Physiological sensors cannot effectively drive mitigation alone. In D. D. Schmorrow, K. M. Stanney, & L. M. Reeves (Eds.), *Foundations of Augmented Cognition, 2nd ed.* (pp. 193-200). Arlington, Virginia: Strategic Analysis Inc.

Hamilton, P. & Tompkins, W.J. (1986). Quantitative investigation of QRS detection rulse using the MIT/BIH arrhythmia database. *IEEE Transactions on Biomedical Engineering, 33 (12),* pp. 1157-1165.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), *Human mental workload.* Amsterdam: North Holland Press.

He, P., Wilson, G. and Russell, C. (2004). Removal of ocular artifacts from electo-encephalogram by adaptive filtering. *Medical and Biological Engineering & Computing, 42,* 407-412.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

He, P., Wilson, G., Russell, C. & Gerschutz, M. (2007). Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data. *Medical and Biological Engineering & Computing, 45,* pp. 495-503.

Hser, Y. -I., & Wickens, T. D. (1989). The effects of the spacing of test trials and study trials in paired-association learning. *Educational Psychology, 9*, 99-120.

Huang, Y., Erdogmus, D., Pavel, M., Mathan, S., & Hild II, K. E. (2011). A framework for rapid visual image search using single trial brain evoked responses. *Neurocomputing, 74*, 2041-2051.

Jasper, H. H. (1958). Report of the Committee on Methods of Clinical Examination. *Electroencephalography and Clinical Neurophysiology, 10*, 370-375.

Kong, X. & Wilson, G.F. (1998). A new EOG-based eyeblink detection algorithm. *Behavior Research Methods, Instruments & Computers, 30 (4),* pp. 713-719.

Krulewitz, J. E., Warm, J. S., & Wohl, T. H. (1975). Effects of shifts in the rate of repetitive stimulation on sustained attention. *Perception and Psychophysics, 18*, 245-249.

Matthews, M. L. (1986). The influence of visual workload history on visual performance. *Human Factors, 28*, 623-632.

McFarland, D. J., Sarnacki, W. A., & Wolpaw, J. R. (2011). Should the parameters of a BCI translation algorithm be continually adapted? *Journal of Neuroscience Methods, 199* (1), 103-107.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors, 33* (1), 17-33.

Pan, J. & Tompkins, W.J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering, 32 (3),* pp. 230-236.

Parasuraman, R., & Wilson, G. F. (2008). Putting the brain to work: neuroergonomics past, present, and future. *Human Factors, 50* (3), 468-474.

Parasuraman, R., Cosenzo, K. A., & De Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: effects on change detection, situation awareness, and mental workload. *Military Psychology, 21* (2), 270-297.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology, 40*, 187-195.

Prinzel III, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *International Journal of Aviation Psychology, 10*, 393-410.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219.

Ros, T., Munneke, M. A., Ruge, D., Gruzelier, J. H., & Rothwell, J. C. (2010). Endogenous control of waking brain rhythms induces neuroplasticity in humans. *European Journal of Neuroscience, 31*, 770-778.

Rouse, W. B. (1976). Adaptive allocation of decision making responsibility between supervisor and computer. In T. B. Sheridan, & G. Johannsen (Eds.), *Monitoring behavior and supervisory control* (pp. 295-306). New York: Plenum Press.

Scerbo, M. (2007). Adaptive automation. In R. Parasuraman, & M. Rizzo (Eds.), *Neuroergonomics: The brain at work* (pp. 238–252). New York: Oxford University Press.

Schmidt, R., Wilson, G. F., Funke, M. A., Davis, I. M., & Caldwell, J. A. (2010). Assessment and classification of cognitive decrements associated with high workload and extended work periods in a UAV setting. *AFRL Technical Report* (AFRL-RH-WP-TP-2010-0015).

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors, 43* (3), 366-380.

Stevens, R., Galloway, T., Berka, C., & Sprang, M. (2009) Can neurophysiologic synchronies provide a platform for adapting team performance? *Lecture Notes in Computer Science*, *5638*, 658-667.

Steriade, M. (2001). Active neocortical processes during quiescent sleep. *Archives of Italian Biology, 139*, 37-51.

Stevens, S. S. (1971). Issues in psychophysical measurement. *Psychological Review*, *78*(5), 426.

Wilson, G. F., & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors, 43* (3), 1005-1018.

Wilson, G. F., & Russell, C. A. (2003). Real-time assessment of mental workload using psychophysiological features and artificial neural networks. *Human Factors, 45* (4), 645-653.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology, 113*, 767-791.

Wolpaw, J. R., McFarland, D. J., Neat, G. W., & Forneris, C. A. (1991). An EEG-basd brain-computer interface for cursor control. *Electroencephalography and Clinical Neurophysiology, 78* (3), 252-259.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

Woolley, A. W., Hackman, J. R., Jerde, T. E., Chabris, C. F., Bennett, S. L., & Kosslyn, S. F. (2007). Using brain-based measures to compose teams: How individual capabilities and team collaboration strategies jointly shape performance. *Social Neuroscience*, *2* (2), 96-105.

Zander, T. O., Kothe, C., Jatzev, S., & Gaertner, M. (2010). Enhancing human-computer interaction with input from passive brain-computer interfaces. In D. S. Tan, & A. Nijholt (Eds.), *Brain-computer interfaces* (pp. 181-199). London: Springer-Verlag London Ltd.

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.

LIST OF ABBREVIATIONS AND ACRONYMS

Ag	silver
AgCl	silver chloride
ANN	artificial neural network
ANOVA	analysis of variance
BCI	brain computer interface
ECG	electrocardiography
EEG	electroencephalography
EOG	electrooculography
HEOG	horizontal (eye movement) electrooculography
HSD	Honestly Significant Difference
Hz	Hertz
IBI	inter-beat interval
IIR	infinite impulse response
NASA-TLX	National Aeronautics and Space Administration Task Load Index
OFS	operator functional state
PC	personal computer
“Physio”	physiological activation of aiding
RPA	remotely piloted aircraft
TLX	Task Load Index
VEOG	vertical (eye movement) electrooculography

Distribution A: Approved for public release; distribution unlimited.
88 ABW Cleared 05/13/2013; 88ABW-2013-2274.